# Quantum Separability and Entanglement Detection via Entanglement-Witness Search and Global Optimization

Lawrence M. Ioannou[1] and Benjamin C. Travaglione[1]

[1]*Centre for Quantum Computation, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK*

We focus on determining the separability of an unknown bipartite quantum state $\rho$ by invoking a sufficiently large subset of all possible entanglement witnesses given the expected value of each element of a set of mutually orthogonal observables. We review the concept of an entanglement witness from the geometrical point of view and use this geometry to show that the set of separable states is not a polytope and to characterize the class of entanglement witnesses (observables) that detect entangled states on opposite sides of the set of separable states. All this serves to motivate a classical algorithm which, given the expected values of a subset of an orthogonal basis of observables of an otherwise unknown quantum state, searches for an entanglement witness in the span of the subset of observables. The idea of such an algorithm, which is an efficient reduction of the quantum separability problem to a global optimization problem, was introduced in PRA 70 060303(R), where it was shown to be an improvement on the naive approach for the quantum separability problem (exhaustive search for a decomposition of the given state into a convex combination of separable states). The last section of the paper discusses in more generality such algorithms, which, in our case, assume a subroutine that computes the global maximum of a real function of several variables. Despite this, we anticipate that such algorithms will perform sufficiently well on small instances that they will render a feasible test for separability in some cases of interest (e.g. in 3-by-3 dimensional systems).

## I. INTRODUCTION

Deciding whether a quantum state, be it physical or theoretical, is separable (as opposed to entangled) is a problem of fundamental importance in the field of quantum information processing and is a computationally intractable problem [1].

One way to decide that a state is entangled is to use an entanglement witness (EW) [2, 3]. Much work has been done on entanglement witnesses (EWs) and their utility in investigating the separability of quantum states, e.g. [4, 5]. EWs have been found to be particularly useful for experimentally detecting the entanglement of states of the particular form $p|\psi\rangle\langle\psi| + (1-p)\sigma$, where $|\psi\rangle$ is an entangled state and $\sigma$ is a mixed state close to the maximally mixed state and $0 \le p \le 1$ [6, 7].

We will show that the set of separable states is not a polytope and thus there is no finite set of EWs that can detect every entangled state. This work focusses on the principle of invoking a sufficiently large subset of all possible EWs given the expected values of a set of observables. Section II summarizes some geometric aspects of the set of Hermitian operators and section III reviews the geometry of separable states and entanglement witnesses. The simplest case of a set of expected values giving rise to more than one EW is characterized in section IV and illustrated by visiting the problem of deciding whether a noisy Bell state is entangled. In section V, we apply the above principle to the problem of detecting the entanglement of a completely unknown quantum state and, in section VI, we outline a class of (classical) algorithms that search for an EW that detects the state or conclude that so such EW exists (given the currently available information about the state).

## II. GEOMETRY OF VECTOR SPACE OF HERMITIAN OPERATORS

Let $\mathcal{H}_{M,N}$ denote the set of all Hermitian operators mapping $\mathbb{C}^M \otimes \mathbb{C}^N$ to itself. This vector space is endowed with the Hilbert-Schmidt inner product $\langle X, Y\rangle \equiv \mathrm{tr}(AB)$, which induces the corresponding norm $||X|| \equiv \sqrt{\mathrm{tr}(X^2)}$ and distance measure $||X-Y||$. By fixing an orthogonal Hermitian basis for $\mathcal{H}_{M,N}$, the elements of $\mathcal{H}_{M,N}$ are in one-to-one correspondence with the elements of the real Euclidean space $\mathbb{R}^{M^2N^2}$. Let $\mathcal{B} = \{X_i : i = 0, 1, \ldots, M^2N^2-1\}$ be an orthonormal, Hermitian basis for $\mathbb{H}_{M,N}$, where $X_0 \equiv \frac{1}{\sqrt{MN}} I$. For concreteness, we can assume that the elements of $\mathcal{B}$ are tensor-products of the (suitably normalized) canonical generators of SU(M) and SU(N), given e.g. in [8]. Note $\mathrm{tr}(X_i) = 0$ for all $i > 0$. Define $v : \mathbb{H}_{M,N} \to \mathbb{R}^{M^2N^2-1}$ as

$$v(A) := \begin{bmatrix} \mathrm{tr}(X_1 A) \\ \mathrm{tr}(X_2 A) \\ \vdots \\ \mathrm{tr}(X_{M^2N^2-1} A) \end{bmatrix}. \tag{1}$$

Via the mapping $v$, the set of separable states $\mathcal{S}_{M,N}$ can be viewed as a full-dimensional convex subset of $\mathbb{R}^{M^2N^2-1}$

$$\{v(\sigma) \in \mathbb{R}^{M^2N^2-1} : \sigma \in \mathcal{S}_{M,N}\}, \tag{2}$$

which properly contains the origin $v(I_{M,N}) = \overline{0} \in \mathbb{R}^{M^2N^2-1}$ (recall that there is a ball of separable states of nonzero radius centred at the maximally mixed state $I_{M,N}$ [9]).

Most of the definitions in the rest of this section may be found in [10]. If $A \in \mathcal{H}_{M,N}$ and $A \neq 0$ and $a \in \mathbb{R}$, then $\{x \in \mathcal{H}_{M,N} : \ \operatorname{tr}(Ax) \leq a\}$ is called the *halfspace* $H_{A,a}$. The boundary $\{x \in \mathcal{H}_{M,N} : \ \operatorname{tr}(Ax) = a\}$ of $H_{A,a}$ is the *hyperplane* $\pi_{A,a}$ with *normal* $A$. Call two hyperplanes *parallel* if they share the same normal. Let $H^\circ_{A,a}$ denote the interior $H_{A,a} \setminus \pi_{A,a}$ of $H_{A,a}$. Note that $H^\circ_{-A,-a}$ is just the complement of $H_{A,a}$. For example, the density operators of an $M$ by $N$ quantum system lie on the hyperplane $\pi_{I,1}$, where $I$ is the identity operator. Let $\mathcal{D}_{M,N} = \{\rho \in \mathcal{H}_{M,N} : \ \rho \geq 0\} \cap \pi_{I,1}$ denote the density operators.

The intersection of finitely many halfspaces is called a *polyhedron*. Every polyhedron is a convex set. Let $D$ be a polyhedron. A set $F \subseteq D$ is a *face* of $D$ if there exists a halfspace $H_{A,a}$ containing $D$ such that $F = D \cap \pi_{A,a}$. If $v$ is a point in $D$ such that the set $\{v\}$ is a face of $D$, then $v$ is a *vertex* of $D$. A *facet* of $D$ is a nonempty face of $D$ having dimension one less than the dimension of $D$. A polyhedron that is contained in a hyperball $\{x \in \mathcal{H}_{M,N} : \operatorname{tr}(x^2) = R^2\}$ of finite radius $R$ is a *polytope*.

## III. SEPARABLE STATES AND ENTANGLEMENT WITNESSES

The set of bipartite separable quantum states $\mathcal{S}_{M,N}$ in $\mathcal{H}_{M,N}$ is defined as the convex hull of the separable pure states $\{|\alpha\rangle\langle\alpha| \otimes |\beta\rangle\langle\beta| \in \mathcal{H}_{M,N}\}_{\alpha,\beta}$, where $|\alpha\rangle$ is a unit vector in $\mathbb{C}^M$ and $|\beta\rangle$ is a unit vector in $\mathbb{C}^N$. Let $\mathcal{E}_{M,N} = \mathcal{D}_{M,N} \setminus \mathcal{S}_{M,N}$ be the set of entangled states. For each entangled state $\rho$ there exists a halfspace $H_{A,a}$ whose interior $H^\circ_{A,a}$ contains $\rho$ but contains no member of $\mathcal{S}_{M,N}$ [2]. Call $A \in \mathcal{H}_{M,N}$ an *entanglement witness* [3] if for some $a \in \mathbb{R}$

$$\mathcal{S}_{M,N} \cap H^\circ_{A,a} = \varnothing \quad \text{and} \quad \mathcal{E}_{M,N} \cap H^\circ_{A,a} \neq \varnothing. \tag{3}$$

Entanglement witnesses $A$ with $a = 0$ in (3) correspond to the conventional definition of "entanglement witness" found in the literature, e.g. [6].

Entanglement witnesses can be used to determine that a physical quantum state is entangled. Suppose $A$ is an EW as in (3) and that a state $\rho$ that is produced in the lab is not known to be separable. If sufficiently many copies of $\rho$ may be produced, then repeatedly measuring the observable $A$ of $\rho$ gives a good estimate of the *expected value of A*

$$\langle A\rangle_\rho := \operatorname{tr}(A\rho)$$

which, if less than $a$, indicates that $\rho \in H^\circ_{A,a}$ and hence that $\rho$ is entangled. Otherwise, if $\langle A\rangle_\rho \geq a$, then $\rho$ may be entangled or separable. The best value of $a$ to use in (3) is $a^* = \min_{|\psi\rangle\langle\psi| \in \mathcal{S}_{M,N}} \{\langle\psi| A |\psi\rangle\}$ since, with this value of $a$, the hyperplane $\pi_{A,a}$ is tangent to $\mathcal{S}_{M,N}$ and thus the volume of entangled states that can be detected by measuring observable $A$ is maximized. With this in mind, define

$$a^*(A) := \min_{|\psi\rangle\langle\psi| \in \mathcal{S}_{M,N}} \{\langle\psi| A |\psi\rangle\}$$

if $A$ is an EW.

Detection of the entanglement of reproducible physical states in the lab would be straightforward if there were a relatively small number $K$ of EWs $A_i$ such that $\mathcal{E}_{M,N}$ is contained in

$$\bigcup_{i=1}^{K} H_{A_i,a_i},$$

where $a_i := a^*(A_i)$. This would imply that $\mathcal{S}_{M,N}$ is

$$\bigcap_{i=1}^{K} H_{-A_i,-a_i},$$

that is, that $\mathcal{S}_{M,N}$ is the intersection of finitely many halfspaces. Invoking the isomorphism between $\mathcal{H}_{M,N}$ and $\mathbb{R}^{M^2N^2}$, this says that $\mathcal{S}_{M,N}$ is a polytope in $\mathbb{R}^{M^2N^2-1}$. Minkowski's theorem [10] says that every polytope in $\mathbb{R}^n$ is the convex hull of its *finitely many* vertices (extreme points). Recall that an extreme point of a convex set is one that cannot be written as a nontrivial convex combination of other elements of the set. To show that $\mathcal{S}_{M,N}$ is not a polytope, it suffices to show that it has infinitely many extreme points. The extreme points of $\mathcal{S}_{M,N}$ are precisely the product states, as we now remind ourselves (see also [11]): A mixed state is not extreme, by definition. Conversely, we have that

$$|\psi\rangle\langle\psi| = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{4}$$

implies

$$1 = \sum_i p_i \,\langle\psi|\, |\psi_i\rangle\langle\psi_i| \,|\psi\rangle = \sum_i p_i |\,\langle\psi_i|\psi\rangle\,|^2, \tag{5}$$

which implies that $|\,\langle\psi_i|\psi\rangle\,| = 1$ for all $i$; thus, a pure state is extreme. Since $\mathcal{S}_{M,N}$ has infinitely many pure product states, we have the following fact, which settles a problem mentioned in [12].

**Fact 1.** *$\mathcal{S}_{M,N}$ is not a polytope in $\mathbb{R}^{M^2N^2-1}$.*

## IV. AMBIDEXTROUS ENTANGLEMENT WITNESSES

Suppose that $A$ is not an entanglement witness but that $-A$ is. In this case, an estimate of $\operatorname{tr}(A\rho)$ is just as useful in testing whether $\rho$ is entangled. We extend the definition of "entanglement witness" to reflect this fact: Call $A \in \mathbb{H}_{M,N}$ a *left (entanglement) witness* if (3) holds for some $a \in \mathbb{R}$, and a *right (entanglement) witness* if

$$\mathcal{S}_{M,N} \cap H^\circ_{-A,-b} = \varnothing \quad \text{and} \quad \mathcal{E}_{M,N} \cap H^\circ_{-A,-b} \neq \varnothing \tag{6}$$

for some $b \in \mathbb{R}$. As well, for $A$ a right witness, define

$$b^*(A) := \max_{|\psi\rangle\langle\psi| \in \mathcal{S}_{M,N}} \{\langle\psi| A |\psi\rangle\}.$$

Note that $A$ is a left witness if and only if $-A$ is a right witness.

The operator $A \in \mathbb{H}_{M,N}$ defines the family $\{\pi_{A,a}\}_{a\in\mathbb{R}}$ of parallel hyperplanes in $\mathbb{R}^{M^2N^2}$. Consider the hyperplane $\pi_A := \pi_{A,\frac{\mathrm{tr}(A)}{MN}}$ which cuts through $\mathcal{S}_{M,N}$ at the maximally mixed state $I_{MN}$. When can $\pi_A$ be shifted parallel to its normal so that it separates $\mathcal{S}_{M,N}$ from some entangled states? If $A$ is *both* a left and right witness, then $\pi_A$ can be shifted either in the positive or negative directions of the normal. In this case, the two parallel hyperplanes $\pi_{A,a^*(A)}$ and $\pi_{A,b^*(A)}$ sandwich $\mathcal{S}_{M,N}$ with some entangled states outside of the *sandwich*, which we will denote by $W(A) := H_{-A,-a^*(A)} \cap H_{-A,-b^*(A)}$.

**Definition 1 (Ambidextrous entanglement witness).** An operator $A \in \mathbb{H}_{M,N}$ is an *ambidextrous (entanglement) witness* if it is both a left witness and a right witness.

If $A$ is an ambidextrous witness, then $\rho$ is entangled if $\langle A\rangle_\rho < a^*(A)$ *or* if $\langle A\rangle_\rho > b^*(A)$. We can further define a *left-handed* witness to be an entanglement witness that is left but not right. Say that two entangled states $\rho_1$ and $\rho_2$ are *on opposite sides of* $\mathcal{S}_{M,N}$ if there does not exist a halfspace $H_{A,a}$ such that $H^\circ_{A,a}$ contains $\rho_1$ and $\rho_2$ but contains no separable states. Ambidextrous witnesses have the potential advantage over conventional (left-handed) entanglement witnesses that they can detect entangled states on opposite sides of $\mathcal{S}_{M,N}$ with the *same* physical measurement.

Entanglement witnesses can be simply characterized by their spectral decomposition. In the following, suppose $A \in \mathbb{H}_{M,N}$ has spectral decomposition $A = \sum_{i=0}^{MN-1} \lambda_i |\lambda_i\rangle\langle\lambda_i|$ with $\lambda_0 \le \lambda_1 \le \ldots \le \lambda_{MN-1}$.

**Fact 2.** *The operator $A$ is a left witness if and only if there exists $k \in [0,1,\ldots,MN-2]$ such that span$(\{|\lambda_0\rangle, |\lambda_1\rangle, \ldots, |\lambda_k\rangle\})$ contains no separable pure states and $\lambda_{k+1} > \lambda_k$.*

*Proof.* Suppose first that there exists no such $k$. Then $|\lambda_0\rangle$ is, without loss of generality, a separable pure state (because the eigenspace corresponding to $\lambda_0$ must contain a product state), so $A$ cannot be a left witness. To prove the converse, suppose that such a $k$ does exist and that $\lambda_{k+1} > \lambda_k$. Define the real function $f(\sigma) := \mathrm{tr}(A\sigma)$ on $\mathcal{S}_{M,N}$. Since span$(\{|\lambda_0\rangle, |\lambda_1\rangle, \ldots, |\lambda_k\rangle\})$ contains no separable states and $\lambda_{k+1} > \lambda_k$, the function satisfies $f(\sigma) > \lambda_0$. Since the set of separable states is compact, there exists a separable state $\sigma'$ that minimizes $f(\sigma)$. Thus, setting $a := f(\sigma')$ gives $\mathcal{S}_{M,N} \cap H^\circ_{A,a} = \varnothing$. As well, $\mathcal{E}_{M,N} \cap H^\circ_{A,a} \neq \varnothing$ since $\mathrm{tr}(A|\lambda_0\rangle\langle\lambda_0|) = \lambda_0 < a$, and so $A$ is a left witness. □

**Theorem 3.** *The operator $A$ is a left or right entanglement witness if and only if (i) there exists $k \in [0,1,\ldots,MN-2]$ such that span$\{|\lambda_0\rangle, |\lambda_1\rangle, \ldots, |\lambda_k\rangle\}$ contains no separable pure states and $\lambda_{k+1} > \lambda_k$, or (ii) there exists $l \in [1,2,\ldots,MN-1]$ such that span$\{|\lambda_l\rangle, |\lambda_{l+1}\rangle, \ldots, |\lambda_{MN-1}\rangle\}$ contains no separable pure states and $\lambda_l > \lambda_{l-1}$.*

Theorem 3 immediately gives a method for identifying and constructing entanglement witnesses.

**Definition 2 (Partial Product Basis, Unextendible Product Basis [13]).** A *partial product basis* of $\mathbb{C}^M \otimes \mathbb{C}^N$ is a set $S$ of mutually orthonormal pure product states spanning a proper subspace of $\mathbb{C}^M \otimes \mathbb{C}^N$. An *unextendible product basis* of $\mathbb{C}^M \otimes \mathbb{C}^N$ is a partial product basis $S$ of $\mathbb{C}^M \otimes \mathbb{C}^N$ whose complementary subspace $(\mathrm{span}S)^\perp$ contains no product state.

We can use unextendible product bases to construct ambidextrous witnesses. Suppose $B$ is an unextendible product basis of $\mathbb{C}^M \otimes \mathbb{C}^N$, and let $B'$ be disjoint from $B$ such that $B \cup B'$ is an orthonormal basis of $\mathbb{C}^M \otimes \mathbb{C}^N$. One possibility is the left witness defined by $A'$ as

$$A' = -\sum_{|\lambda\rangle \in B'} |\lambda\rangle\langle\lambda| \tag{7}$$

As well, we could split $B'$ into $B'_L$ and $B'_R$ and define an ambidextrous witness $A''$ as

$$A'' = -\sum_{|\lambda_L\rangle \in B'_L} |\lambda_L\rangle\langle\lambda_L| + \sum_{|\lambda_R\rangle \in B'_R} |\lambda_R\rangle\langle\lambda_R|. \tag{8}$$

Another thing to realize is that span$B$ may contain an entangled pure state, which can be pulled out and put into a $(+1)$-eigenvalue eigenspace of $A'$. Depending on $B$ (and the dimensions $M$, $N$), there may be several mutually orthogonal pure entangled states in span$B$ whose span contains no product state; let $B''$ be a set of such pure states. Define the ambidextrous witness as

$$A''' = -\sum_{|\lambda\rangle \in B'} |\lambda\rangle\langle\lambda| + \sum_{|\lambda\rangle \in B''} |\lambda\rangle\langle\lambda|. \tag{9}$$

This suggests the following problem, related to the combinatorial [14] problem of finding unextendible product bases: Given $M$ and $N$, find all orthonormal bases $B$ for $\mathbb{C}^M \otimes \mathbb{C}^N$ such that

- $B$ is the disjoint union of $\Lambda_L$, $\tilde{B}$, $\Lambda_R$,
- span$\Lambda_L$ and span$\Lambda_R$ contain no product state,
- span$(\Lambda_L \cup \Lambda_R)$ contains a product state, and
- $\min\{|\Lambda_L|, |\Lambda_R|\}$ is maximal.

Such bases may give "optimal" ambidextrous witnesses, which detect the largest volume of entangled states on opposite sides of $\mathcal{S}_{M,N}$.

The functions $a^*$ and $b^*$ are difficult to compute [29]. Thus a criticism of constructing witnesses via the spectral decomposition is that even if you can construct the

corresponding physical observables, you still have to perform a difficult computation to make them useful. However, most experimental applications of entanglement witnesses are in very low dimensions, where computing $a^*$ and $b^*$ deterministically is not a problem – it may even be done analytically, as in the example at the end of this section.

Ambidextrous witnesses represent the simplest case of the principle of invoking as many (left) entanglement witnesses as possible given the expected values of each element of a set $X$ of linearly independent observables, that is, the case $|X| = 1$. In the Section V, we see how this principle generalizes to $|X| > 1$.

A simple illustration of how AEWs may be used involves detecting and distinguishing noisy Bell states. Define the four Bell states in $\mathbb{C}^2 \otimes \mathbb{C}^2$:

$$\begin{aligned} \left|\psi^{\pm}\right\rangle &:= \left(|00\rangle \pm |11\rangle\right)/\sqrt{2} \\ \left|\phi^{\pm}\right\rangle &:= \left(|01\rangle \pm |10\rangle\right)/\sqrt{2}. \end{aligned}$$

It is straightforward to show that the Bell states are, pairwise, on opposite sides of $\mathcal{S}_{2,2}$. Suppose a left entanglement witness $W$, with $a^*(W) = 0$, detects $|\psi^+\rangle$ and $|\phi^+\rangle$. Without loss of generality, $W$ can be written in the Bell basis $\{|\psi^+\rangle, |\phi^+\rangle, \ldots\}$ as

$$W = \begin{bmatrix} -\epsilon_1 & a+bi & \times & \times \\ a-bi & -\epsilon_2 & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}, \qquad (10)$$

for $\epsilon_1$ and $\epsilon_2$ both positive. But the states $|s^{\pm}\rangle \equiv \frac{1}{\sqrt{2}}(|\psi^+\rangle \pm |\phi^+\rangle)$ are separable. Requiring $\langle s^+|\, W\, |s^+\rangle \geq 0$ gives $2a \geq \epsilon_1 + \epsilon_2$ and requiring $\langle s^-|\, W\, |s^-\rangle \geq 0$ gives $2a \leq -\epsilon_1 - \epsilon_2$, which, together, give a contradiction. Similar arguments hold for the other pairs of Bell states.

Define the operators

$$\begin{aligned} A_\psi &:= -|\psi^-\rangle\langle\psi^-| + |\psi^+\rangle\langle\psi^+| \\ A_\phi &:= -|\phi^-\rangle\langle\phi^-| + |\phi^+\rangle\langle\phi^+|. \end{aligned}$$

Both $A_\psi$ and $A_\phi$ are easily seen to be AEWs. It is also straightforward to compute the values

$$a^*(A_\psi) = a^*(A_\phi) = -1/2$$

and

$$b^*(A_\psi) = b^*(A_\phi) = +1/2.$$

Suppose that there is a source that repeatedly emits the same noisy Bell state $\rho$ and that we want to decide whether $\rho$ is entangled. Define the Pauli operators:

$$\begin{aligned} \sigma_0 &:= \tfrac{1}{\sqrt{2}}(|0\rangle\langle 0| + |1\rangle\langle 1|) \\ \sigma_1 &:= \tfrac{1}{\sqrt{2}}(|0\rangle\langle 1| + |1\rangle\langle 0|) \\ \sigma_2 &:= -\tfrac{i}{\sqrt{2}}(|0\rangle\langle 1| - |1\rangle\langle 0|) \\ \sigma_3 &:= \tfrac{1}{\sqrt{2}}(|0\rangle\langle 0| - |1\rangle\langle 1|)\ , \end{aligned}$$

where $\{|0\rangle, |1\rangle\}$ is the standard orthonormal basis for $\mathbb{C}^2$. Noting that

$$\begin{aligned} A_\psi &= \sigma_1 \otimes \sigma_1 - \sigma_2 \otimes \sigma_2 \\ A_\phi &= \sigma_1 \otimes \sigma_1 + \sigma_2 \otimes \sigma_2, \end{aligned}$$

measuring the expected value of the two observables $\sigma_1 \otimes \sigma_1$ and $\sigma_2 \otimes \sigma_2$ may be sufficient to decide that $\rho$ is entangled because $\rho \in \mathcal{E}_{2,2}$ if one of the following four inequalities is true:

$$\begin{aligned} \langle \sigma_1 \otimes \sigma_1 \rangle_\rho \pm \langle \sigma_2 \otimes \sigma_2 \rangle_\rho &> 1/2 \\ \langle \sigma_1 \otimes \sigma_1 \rangle_\rho \pm \langle \sigma_2 \otimes \sigma_2 \rangle_\rho &< -1/2. \end{aligned} \qquad (11)$$

If the noise is known to be of a particular form, then we can also determine *which* noisy Bell state was being produced. Let $|B\rangle$ be a Bell state. Suppose $\rho$ is known to be of the form $p|B\rangle\langle B| + (1-p)\sigma$ for some $\sigma$ inside both sandwiches $W(A_\psi)$ and $W(A_\phi)$. With $\sigma$ so defined, one of the four inequalities (11) holds only if exactly one of them holds, so that $|B\rangle$ is determined by which inequality is satisfied. We remark that, if $\sigma$ and $|B\rangle$ are known, knowledge of the expected value of any single observable $A$ may allow one to compute $p$ and hence an upper bound on the $l_2$ distance between $\rho$ and the maximally mixed state $I/4$. This distance may be enough information to conclude that $\rho$ is separable by checking if $\rho$ is inside the largest separable ball centred at $I/4$ [9].

## V. DETECTING ENTANGLEMENT OF AN UNKNOWN STATE USING PARTIAL INFORMATION

We now consider the task of trying to decide whether a completely unknown physical state $\rho$, of which many copies are available, is entangled. For simplicity, we restrict to $\rho \in \mathcal{H}_{2,2}$ but the discussion can be applied to a bipartite system of any dimension, replacing Pauli operators with canonical generators of SU(M) and SU(N) (or any linearly independent Hermitian product basis). For such $\rho$, this problem has already been addressed in [15], where the so-called “structural physical approximation of an unphysical map” [16] was used to implement the Peres-Horodecki positive partial transpose (PPT) test [2, 17]. While the structural physical approximation is experimentally viable in principle, it is currently very difficult to do so. Thus, the easiest way to test for entanglement at present is to perform “state tomography” in order to get good estimates of 15 real parameters that define $\rho$, then reconstruct the density matrix for $\rho$ and carry out the PPT test [17] on this matrix.

An experimentalist has many choices of which 15 parameters to estimate: the expectations of any 15 linearly independent observables qualify, as do the probability distributions of any 5 mutually unbiased (four-outcome) measurements [18, 19]. Whatever 15 parameters are chosen, we assume that the basic tool of the experimentalist is the ability to perform local two-outcome measurements

on each qubit, e.g. measuring $\sigma_1$ on the first qubit and $\sigma_2$ on the second. Under this assumption, the scenario where the two qubits of $\rho$ are far apart is easily handled if classical communication is allowed between the two labs. We further assume, for simplicity, that the set of these local two-outcome measurements is the set of Pauli operators $\{\sigma_i\}_{i=0,1,2,3}$ (defined on page 4). If $\sigma_i$ is measured on the first qubit and $\sigma_j$ on the second, repeating this procedure on many copies of $\rho$ gives good estimations of the three expectations $\langle\sigma_i \otimes \sigma_0\rangle$, $\langle\sigma_0 \otimes \sigma_j\rangle$, and $\langle\sigma_i \otimes \sigma_j\rangle$ (where the subscript "$\rho$" is omitted for readability). Let us call this procedure *measuring* $\sigma_i\sigma_j$.

Suppose the experimentalist sets out to solve our problem and begins the data collection by measuring $\sigma_1\sigma_1$ and then $\sigma_2\sigma_2$. Even though only 6 of the 15 independent parameters defining $\rho$ have been found, the example in the previous section shows that $\rho$ is entangled if one of the four inequalities (11) is true. It is straightforward to show that if none of these inequalities is true, then no entanglement witness in the span of $\{\sigma_1 \otimes \sigma_1, \sigma_2 \otimes \sigma_2\}$ can detect $\rho$ if it is entangled [30]. However, there may be an entanglement witness in the span of

$$\{\sigma_0 \otimes \sigma_1, \sigma_0 \otimes \sigma_2, \sigma_1 \otimes \sigma_1, \sigma_2 \otimes \sigma_2, \sigma_1 \otimes \sigma_0, \sigma_1 \otimes \sigma_0\}$$

that does detect $\rho$.

More generally, at any stage of the data-gathering process, if we have the set of expectations $\{\langle\sigma_i \otimes \sigma_j\rangle : (i, j) \in T\}$, then $\rho$ is entangled if there is an entanglement witness in the span of $\{\sigma_i \otimes \sigma_j : (i, j) \in T\}$ that detects $\rho$ ($T \subset \{(k, l) : k, l \in \{0, 1, 2, 3\}\} \setminus (0, 0)$). If the experimentalist has access to a computer program that can quickly discover such an entanglement witness (if it exists), then the data-gathering process can be terminated early and no more qubits have to be used to decide that $\rho$ is entangled. The algorithms described in the next section are just such programs. To see this, note that the projection $\overline{\mathcal{S}}_{2,2}$ of $\mathcal{S}_{2,2}$ onto span$\{\sigma_i \otimes \sigma_j : (i, j) \in T\}$ is a full-dimensional convex subset of $\mathbb{R}^{|T|}$, and the projection $\overline{\rho}$ of $\rho$ onto span$\{\sigma_i \otimes \sigma_j : (i, j) \in T\}$ is a point in $\mathbb{R}^{|T|}$ such that $\overline{\rho} \notin \overline{\mathcal{S}}_{2,2}$ if and only if there is an entanglement witness in the span of $\{\sigma_i \otimes \sigma_j : (i, j) \in T\}$ that detects $\rho$. Since the following algorithms can be applied to any full-dimensional convex set (satisfying certain conditions), we can apply them to $\overline{\mathcal{S}}_{2,2}$.

We view any such algorithm as an extra tool that an experimentalist can use to facilitate entanglement detection and minimize the number of copies of $\rho$ that must be measured – essentially, trading classical resources for quantum resources. As we saw in the case of constructing ambidextrous witnesses, the primary classical resource required to invoke a sufficiently large subset of all such entanglement witnesses is a subroutine for computing the function $b^*$ (equivalently, $a^*$).

## VI. ALGORITHMS FOR FINDING ENTANGLEMENT WITNESSES BASED ON GLOBAL OPTIMIZATION

Assume that $\rho \in \mathcal{D}_{M,N}$ is a state whose separability is unknown. We can handle two scenarios – one experimental, as described above, and the other theoretical. In the theoretical scenarios, we assume that we know the density matrix for $\rho$; this corresponds to having gathered all $M^2N^2 - 1$ independent expected values in the experimental scenario. Since the algorithms find an entanglement witness when $\rho \in \mathcal{E}_{M,N}$, they could also be applied when $\rho$ is known to be entangled but an entanglement witness for $\rho$ is desired (though one may want to apply the entanglement witness optimization procedure [4] to the result of the algorithm, as these algorithms do not necessarily output optimal entanglement witnesses).

Let $j$ be the number of nontrivial expected values of $\rho$ that are known, $2 \leq j \leq M^2N^2 - 1$; that is, (without loss) assume we know the expected values of the elements of $\mathcal{B}' = \{X_1, X_2, \ldots, X_j\}$. The algorithms either find an entanglement witness in span($\mathcal{B}'$) for $\rho$, or conclude that no such witness exists. For any $Y \in \mathbb{H}_{M,N}$ with $Y = \sum_{i=0}^{M^2N^2-1} y_i X_i$, let $\overline{Y}$ be the $j$-dimensional vector of the real numbers $y_i$ for $i = 1, 2, \ldots, j$. Define

$$\overline{\mathcal{S}}_{M,N} = \{\overline{\sigma} : \sigma \in \mathcal{S}_{M,N}\}. \tag{12}$$

Note $\overline{\mathcal{S}}_{M,N}$ is a full-dimensional convex set in $\mathbb{R}^j$, properly containing the origin (since $\overline{I_{M,N}}$ is the zero-vector in $\mathbb{R}_j$).

Let $K$ be a full-dimensional convex subset of $\mathbb{R}^n$ which contains a ball of finite nonzero radius centred at the origin and is contained in a ball of finite radius $R$. The algorithms are general and can be used to decide whether a hyperplane exists which separates a given point $p$ from any given $K$ satisfying these properties. Thus, the clearest way to describe how the algorithms work is to use the application-neutral notation of convex analysis. For $x \in \mathbb{R}^n$ and $\delta > 0$, let $B(x, \delta) := \{y \in \mathbb{R}^n : ||x - y|| \leq \delta\}$. For a convex subset $K \subset \mathbb{R}^n$, let $S(K, \delta) := \cup_{x \in K} B(x, \delta)$ and $S(K, -\delta) := \{x : B(x, \delta) \subseteq K\}$. Define the following convex body problems [20]:

**Definition 3 (Weak separation problem for $K$ (WSEP($K$))).** Given a rational vector $p \in \mathbb{R}^n$ and rational $\delta > 0$, either

- assert $p \in S(K, \delta)$, or
- find a rational vector $c \in \mathbb{R}^n$ with $||c||_\infty = 1$ such that $c^T x < c^T p$ for every $x \in K$ [31].

**Definition 4 (Weak optimization problem for $K$ (WOPT($K$))).** Given a rational vector $c \in \mathbb{R}^n$ and rational $\epsilon > 0$, either

- find a rational vector $y \in \mathbb{R}^n$ such that $y \in S(K, \epsilon)$ and $c^T x \leq c^T y + \epsilon$ for every $x \in K$; or
- assert that $S(K, -\epsilon)$ is empty [32].

By taking $\delta$ and $\epsilon$ to be zero, we implicitly define the corresponding *strong* separation (SSEP) and *strong* optimization (SOPT) problems. Note that by taking $K$ to be $\overline{\mathcal{S}}_{M,N}$ and $p$ to be $\overline{\rho}$ (for some state $\rho \in \mathcal{D}_{M,N}$), SSEP($K$) corresponds to the problem of finding an entanglement witness for $\rho$ (or deciding that one does not exist in the span of $\mathcal{B}'$) [33]; and by further taking $c$ to be $\overline{A}$, for some $A \in \mathbb{H}_{M,N}$, SOPT($K$) corresponds to the problem of computing $b^*(A)$ (actually something at least as hard, since SOPT($K$) asks for the *maximizer* of $c^T x$, over $x \in K$, rather than just the *maximum*).

We describe *oracle-polynomial-time* algorithms for WSEP($K$), with respect to an oracle for WOPT($K$); that is, assuming each call to the oracle is assigned unit complexity cost, the algorithms have running time in $O(\text{poly}(n, \log(R/\delta)))$. We will use "$\mathcal{O}$" to denote oracles (black-boxed subroutines) for problems, indicating which problem via a subscript, e.g. $\mathcal{O}_{\text{SOPT}(K)}$. In what follows, so as not to obfuscate the main idea of the algorithms, we ignore the weakness of the separation and optimization problems; that is, we assume we are solving SSEP($K$) with an oracle for SOPT($K$).

There are at least two ways to reduce SSEP($K$) to SOPT($K$). The first method was covered in [21, 22]; the second method, which we give below, is well known and may be found in the synthesis of Lemma 4.4.2 and Theorem 4.2.2 in [20]. For $y \in \mathbb{R}^n$ and $b \in \mathbb{R}$, define the hyperplane $\pi_{y,b} \equiv \{x \in \mathbb{R}^n : y^T x = b\}$.

**Definition 5 (Polar of $K$).** The *polar* $K^\star$ of a full-dimensional convex set $K \subset \mathbb{R}^n$ that contains the origin is defined as

$$K^\star := \{c \in \mathbb{R}^n : c^T x \le 1 \ \ \forall x \in K\}. \tag{13}$$

If $c \in K^\star$, then the plane $\pi_{c,1}$ separates $p \in \mathbb{R}^n$ from $K$ when $c^T p > 1$.

**Definition 6 (Feasibility problem for $K'$ (FEAS($K'$))).** Given a convex set $K' \subset \mathbf{R}^n$, either

- find a point $k' \in K'$, or
- assert that $K'$ is empty.

Thus, the separation problem for $p$ is equivalent to the feasibility problem for $Q_p$, defined as

$$Q_p := K^\star \cap \{c : p^T c \ge 1\}. \tag{14}$$

As outlined in the next section, to solve the feasibility problem for any $K'$, it suffices to have a separation routine for $K'$. Because we can easily build a separation routine $\mathcal{O}_{\text{SSEP}(Q_p)}$ for $Q_p$ out of $\mathcal{O}_{\text{SSEP}(K^\star)}$, it suffices to have a separation routine $\mathcal{O}_{\text{SSEP}(K^\star)}$ for $K^\star$ in order to solve the feasibility problem for $Q_p$ [34]. Building $\mathcal{O}_{\text{SSEP}(Q_p)}$ out of $\mathcal{O}_{\text{SSEP}(K^\star)}$ is done as follows:

Routine $\mathcal{O}_{\text{SSEP}(Q_p)}(y)$:
CASE: $p^T y < 1$
  RETURN $-p$
ELSE: $p^T y \ge 1$
  CALL $\mathcal{O}_{\text{SSEP}(K^\star)}(y)$
  CASE: $\mathcal{O}_{\text{SSEP}(K^\star)}(y)$ returns separating vector $q$
    RETURN $q$
  ELSE: $\mathcal{O}_{\text{SSEP}(K^\star)}(y)$ asserts $y \in K^\star$
    RETURN "$y \in Q$"

It remains to show that the optimization routine $\mathcal{O}_{\text{SOPT}(K)}$ for $K$ gives a separation routine $\mathcal{O}_{\text{SSEP}(K^\star)}$ for $K^\star$. Suppose $y$ is given to $\mathcal{O}_{\text{SOPT}(K)}$, which returns $k \in K$ such that $y^T x \le y^T k =: b$ for all $x \in K$. If $b \le 1$, then $\mathcal{O}_{\text{SSEP}(K^\star)}$ may assert $y \in K^\star$. Otherwise, $\mathcal{O}_{\text{SSEP}(K^\star)}$ may return $k$, because $\pi_{k,1}$ (and hence $\pi_{k,b}$) separates $y$ from $K^\star$: since $k^T y = b > 1$, it suffices to note that $k^T c = c^T k \le 1$ for all $c \in K^\star$ by the definition of $K^\star$ and the fact that $k \in K$.

The plane $\pi_{k,1}$ is called a *cutting plane*, and, to solve FEAS($K'$) with $\mathcal{O}_{\text{SSEP}(K')}$, we use a *cutting-plane algorithm.* All such algorithms have the same basic structure:

1. Define a (possibly very large) regular bounded convex set $P_0$ which is guaranteed to contain $K'$, such that, for some reasonable definition of "centre", the centre $\omega_0$ of $P_0$ is easily computed. The set $P_0$ is called an *outer approximation* to $K'$. Common choices for $P_0$ are the origin-centred hyperbox, $\{x \in \mathbb{R}^n : -2^L \le x_i \le 2^L, \ 1 \le i \le n\}$ and the origin-centred hyperball, $\{x : x^T x \le 2^L\}$ (where $2^L$ is a trivially large bound).

2. Give the centre $\omega$ of the current outer approximation $P$ to $\mathcal{O}_{\text{SSEP}(K')}$.

3. If $\mathcal{O}_{\text{SSEP}(K')}$ asserts "$\omega \in K'$", then HALT.

4. Otherwise, say $\mathcal{O}_{\text{SSEP}(K')}$ returns the cutting plane $\pi_{c,b}$ such that $K' \subset \{x : c^T x \le b\}$. Update (shrink) the outer approximation $P := P \cap \{x : c^T x \le b'\}$ for some $b' \ge b$; the idea is that the new $P$ has about half the volume of the old $P$ (i.e. usually $\pi_{c,b}$ passes through $\omega$, or nearby it). Possibly perform other computations to further update $P$. Check stopping conditions; if they are met, then HALT. Otherwise, go to step 2.

The difficulty with such algorithms is knowing when to halt in step 4. Generally, the stopping conditions are related to the size of the current outer approximation. Because it is always an approximate (weak) feasibility problem that is solved, the associated accuracy parameter $\delta'$ can be exploited to get a "lower bound" $V$ on the "size" of $K'$, with the understanding that if $K'$ is smaller than this bound, then the algorithm can correctly assert that $S(K', -\delta')$ is empty. Thus the algorithm stops in step 4 when the current outer approximation is smaller than $V$.

Using the above reduction from SSEP($K$) to SOPT($K$), there are a number of polynomial-time convex

feasibility algorithms that can be applied (see [23] for a discussion of all of them). The three most important are the *ellipsoid method*, the *volumetric-centre method*, and the *analytic-centre method*. The latter are more efficient than the ellipsoid algorithm and are very similar to each other in complexity and precision requirements, with the analytic-centre cutting-plane (ACCP) algorithm in [23] having some supposed practical advantages.

We refer to [22] and [23] (and references therein) for a discussion of details, including computer precision requirements, for either of the ACCP algorithms arising from either the reduction in [21] or the well-known one given here.

## VII. CLOSING REMARKS

In terms of attempting to find a practical algorithm for the quantum separability problem, the skeptic notices that such algorithms appear not to offer any advantage over other approaches: instead of having to solve one instance of an NP-hard problem, we now have to solve many! In response to such skepticism, we can, at least, refer to [24], where it is shown that the asymptotic complexity of such algorithms compares well with all other deterministic algorithms (with known worst-case complexity bounds) for the quantum separability problem. There are many algorithms available for optimizing functions and thus computing $\mathrm{WOPT}(\mathcal{S}_{M,N})$, including the semidefinite programming relaxation method of Lasserre [25] (on which, incidentally, one can base a different separability algorithm [26]); Lipschitz optimization [27]; and Hansen's global optimization algorithm using interval analysis [28].

## VIII. ACKNOWLEDGEMENTS

We would like to thank Carolina Moura Alves, Coralia Cartis, and Tom Stace for helpful discussions and assistance. This work was supported by the CESG, NSERC, ORS, RESQ (EU grant IST-2001-37559), and CMI.

---


[1] L. Gurvits, in *Proceedings of the thirty-fifth ACM symposium on Theory of computing* (ACM Press, New York, 2003), pp. 10–19.

[2] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[3] B. M. Terhal, Phys. Lett. A **271**, 319 (2000).

[4] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, Phys. Rev. A **62**, 052310 (2000).

[5] M. Lewenstein, B. Kraus, P. Horodecki, and J. I. Cirac, Phys. Rev. A **63**, 044304 (2001).

[6] O. Gühne, P. Hyllus, D. Bruß, A. Ekert, M. Lewenstein, C. Macchiavello, and A. Sanpera, Phys. Rev. A **66**, 062305 (2002).

[7] M. Barbieri, F. D. Martini, G. D. Nepi, P. Mataloni, G. M. D'Ariano, and C. Macchiavello, *Experimental detection of entanglement with polarized photons* (2003), quant-ph/0307003.

[8] R. T. Thew, K. Nemoto, A. G. White, and W. J. Munro, Phys. Rev. A **66**, 012303 (2002).

[9] L. Gurvits and H. Barnum, Phys. Rev. A **66**, 062311 (2002).

[10] G. L. Nemhauser and L. Wolsey, *Integer and Combinatorial Optimization* (John Wiley and Sons, Chichester, 1988).

[11] P. Horodecki, Phys. Lett. A **232**, 333 (1997).

[12] D. Bruß, J. Math. Phys. **43**, 4237 (2002).

[13] B. M. Terhal, Lin. Alg. Appl. **323**, 61 (2001).

[14] N. Alon and L. Lovász, Journal of Combinatorial Theory, Series A **95**, 169 (2001).

[15] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2002).

[16] P. Horodecki, *From limits of quantum nonlinear operations to multicopy entanglement witnesses and state spectrum estimation* (2001), quant-ph/0111036.

[17] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[18] I. D. Ivanović, J. Phys. A: Math. Gen. **14**, 3241 (1981).

[19] W. K. Wootters and B. D. Fields, Annals of Physics **191**, 363 (1989).

[20] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization* (Springer-Verlag, Berlin, 1988), ISBN 038713624x.

[21] L. M. Ioannou, B. C. Travaglione, D. C. Cheung, and A. K. Ekert, Phys. Rev. A **70**, 060303(R) (2004).

[22] L. M. Ioannou, *Computing finite-dimensional bipartite quantum separability* (2005), PhD thesis, available at http://arXiv.org/abs/cs/0504110.

[23] D. S. Atkinson and P. M. Vaidya, Mathematical Programming **69**, 1 (1995).

[24] L. M. Ioannou, *Deterministic computational complexity of the quantum separability problem* (2006), in preparation, quant-ph/0603199.

[25] J. B. Lasserre, SIAM J. Optim. **11(3)**, 796 (2001).

[26] J. Eisert, P. Hyllus, O. Gühne, and M. Curty, Phys. Rev. A **70**, 062317 (2004).

[27] R. Horst and P. Pardalos, eds., *Handbook of Global Optimization* (Kluwer Academic Publishers, Dordrecht, 1995).

[28] E. Hansen, *Global Optimization Using Interval Analysis* (Marcel Dekker Incorporated, Boston, 1992), ISBN 0824786963.


[29] At least, (see Section VI) $\mathrm{WOPT}(\mathcal{S}_{M,N})$ is an NP-hard problem, because $\mathrm{WSEP}(\mathcal{S}_{M,N})$ is both NP-hard [1, 20] and, as the existence of the algorithms described in Section VI proves, efficiently reducible to $\mathrm{WOPT}(\mathcal{S}_{M,N})$.

[30] To show this, it suffices to find four separable states whose projections onto $\mathrm{span}\{\sigma_1 \otimes \sigma_1, \sigma_2 \otimes \sigma_2\}$ are the four vertices of the square with vertices $(\frac{1}{2}, 0)$, $(0, \frac{1}{2})$, $(-\frac{1}{2}, 0)$, and $(0, -\frac{1}{2})$; such states are $\frac{1}{4}I \pm \frac{1}{2}\sigma_i \otimes \sigma_i$ for $i = 1, 2$. The result then follows from convexity of $\mathcal{S}_{2,2}$.

[31] The $l_\infty$ norm appears here as a technicality, so that $c$ need not be normalized by a possibly irrational multiplier. We will just use the Euclidean norm in what follows and have

$||c|| \approx 1$.

[32] This will never be the case for us, as $\mathcal{S}_{M,N}$ is not empty.

[33] If $p$ arises from some estimation procedure (as in our experimental setting), then there is a hyperbox around $p$ that contains the "actual" point $p^{\checkmark}$; the hyperbox is given by the "error bars" on each coordinate of $p$. From the "error bars" can be computed a $\Delta > 0$ such that $p^{\checkmark} \in B(p, \Delta)$. If the WSEP($K$) algorithm asserts $p \in S(K, \delta)$, then we can only assert that $p^{\checkmark}$ is in $S(K, \delta{+}\Delta)$; otherwise, we can only assert that $c^T x < c^T p^{\checkmark} + \Delta$ for every $x \in K$.

[34] We slightly abuse the oracular "$\mathcal{O}$" notation by using it for both truly oracular (black-boxed) routines and for other (possibly not completely black-boxed) routines.